Comment on *"What longevity predictors should be allowed for when valuing pension-scheme liabilities?"*, a paper presented to the Institute of Actuaries' sessional meeting on 28th September 2009 by Madrigal *et al.*

During the debate a question was raised over the handling of fractional years of exposure. In ¶4.3.1 the authors write that they "weight the contribution of each of the membership records according to its exposure to risk in a year, counting as a full observation when the exposed to risk is equal to one.". This suggests that the authors are using a weighted log-likelihood, ℓ^w, as follows:

$$\ell^w = w(1-d)\log(1-q_x) + wd\log q_x$$

where q_x denotes the one-year probability of death at age x, d is an indicator variable taking the value 0 upon survival and 1 upon death and w is the exposure. The description in ¶4.3.1 suggests that w takes the value 1 where a complete year of exposure applies and some smaller positive value when a complete year is not possible.

If the authors are doing this then they are creating a bias in their model, even if the weights are not quite as described above. They are taking observations which by definition have artificially low observed mortality by virtue of being only observed for part of a year. They are then including them in a model with a smaller weight than proper observations for a complete year (or potential year). All this achieves is reducing the *influence* of the observations with incomplete years, it does not allow for the incompleteness of the years themselves. The effect would be to produce lower mortality rates than should be the case, i.e. the resulting estimates would be biased estimates of the true underlying mortality rates. The extent of any bias would be directly linked to the relative proportions of part-year and whole-year observations. Another point is that the weight w should not be applied when $d = 1$, as a lower number of expected deaths given fractional exposures should take care of itself without weighting.

To illustrate this, consider the extreme situation where the only data you have is of partial years of exposure. Imagine an annualised mortality rate of 0.2 amongst a group of 100 identical individuals. In a complete year we would therefore expect 20 deaths. If we only had half a year's exposure, then there would be on average only 10 deaths: 20 * 0.5 = 10, assuming a uniform distribution of deaths (UDD) throughout the year. However, if records are weighted according to exposure, the estimated annual mortality rate would be 10 * 0.5 / 100 * 0.5 = 0.1. In their admirable aim of trying to include fractional years of exposure, the authors may have inadvertently created a material source of bias in their model.

There are a number of ways to properly allow for fractional exposure in q_x models, including UDD, constant force of mortality or the Balducci assumption. However, none of these fits easily into the chosen GLM framework. The cleanest approach is simply to use survival models as they automatically allow for fractional years of exposure. This is because they model the time to an event, not the number of events occurring.

Stephen Richards
7$^{\text{th}}$ October 2009